CLAIMS

We claim:

1

2

3.

(h)

- 1 A method of generating a natural language understanding (NLU) model for use 1. 2 in a spoken dialog service, the method comprising: 3 (a) using sample utterances, creating at least one hand crafted rule for each 4 call-type defined in a labeling guide; 5 generating and testing a first NLU model using the at least one hand 6 crafted rule and sample utterances; 7 (c) building a second NLU model using the sample utterances as new 8 training data and using the at least one hand crafted rule; 9 testing the performance of the second NLU model using a first batch of (d) 10 labeled data; 11 (e) building a series of NLU models by adding a previous batch of labeled 12 data to training data and using a new batch of labeling data as test data to generate the series of NLU models with training data that increases constantly; 13 14 (f) if not all the labeling data is received, repeat step (e) until all labeling data 15 is received; and 16 (g) after all the training data is received, at least once, building a third NLU 17 model using all the labeling data, wherein the third NLU model is used in generating the 18 spoken dialog service. 2. 1 The method of claim 1, further comprising: 2 applying at least one obvious utterance to the NLU model and checking whether 3 the at least one obvious utterance was correctly classified.
 - 13

performing an improvement performance step on the third NLU model;

The method of claim 1, further comprising:

- 3 (i) adding the test data to the training data and building a fourth NLU
- 4 model; and
- 5 (j) augmenting the fourth NLU model using utterances obtained from a
- 6 customer acceptance test.
- 1 5. The method of claim 4, further comprising ignoring the utterances that are longer
- 2 than a specific threshold.
- 1 6. The method of claim 5, wherein the threshold is fifty words.
- 1 7. The method of claim 1, wherein if the sample utterances are human-human
- 2 utterances, the method further comprises not labeling utterances longer then a threshold
- 3 amount.
- 1 8. The method of claim 7, wherein the threshold is 80 words.
- 1 9. The method of claim 1, wherein some examples in the labeling guide are ignored.
- 1 10. The method of claim 1, wherein the step of building a series of NLU models by
- 2 adding a previous batch of labeled data to training data and using a new batch of labeling
- data as test data to generate the series of NLU models with training data;
- 1 11. A method of generating a natural language understanding (NLU) model for use
- 2 in a spoken dialog service, the method comprising:
- 3 (a) building a first NLU model using samples utterance from a labeling
- 4 guide, hand crafted rules and labeled utterances of available human/human dialogs or
- 5 human/machine dialogs, if available;
- 6 (b) testing the performance of the first NLU model using sample utterances
- 7 in the labeling guide;

8	(c)	building a series of NLU models and evaluating the performance of the
9	series of NLU	models as labeled data becomes available by:

- 10 (i) adding a previous batch of labeled data to training data;
- 11 (ii) using a new batch of labeling data as test data to generate the
- series of NLU models with training data that increases constantly.
- 1 12. The method of claim 11, wherein the step of building a series of NLU models
- 2 further comprises mixing all the data and dividing the mixed data into training data and
- 3 test data.
- 1 13. The method of claim 11, wherein the step of building a series of NLU models
- 2 further comprises keeping a fixed text set and using all other data for training.
- 1 14. The method of claim 11, wherein the step of building a series of NLU models
- 2 further comprises mixing all the data and dividing it into training and text data.
- 1 15. The method of claim 11, wherein one batch containing all the data is received
- 2 after completion of all the data labeling.
- 1 16. A natural language understanding (NLU) model for use in a spoken dialog
- 2 system, the NLU model generated according to a process comprising:
- 3 (a) building a first NLU model using samples utterance from a labeling
- 4 guide, hand crafted rules and labeled utterances of available human/human dialogs or
- 5 human/machine dialogs, if available;
- 6 (b) testing the performance of the first NLU model using sample utterances
- 7 in the labeling guide;
- 8 (c) building a series of NLU models and evaluating the performance of the
- 9 series of NLU models as labeled data becomes available by:
- 10 (i) adding a previous batch of labeled data to training data; and

- 11 (ii) using a new batch of labeling data as test data to generate the
- series of NLU models with training data that increases constantly.
- 1 17. The NLU model of claim 16, wherein the NLU model is generated by a method
- 2 further comprising:
- 3 augmenting the series of NLU models using utterances obtained from a customer
- 4 acceptance test.
- 1 18. The NLU model of claim 17, wherein the NLU model is generated by a method
- 2 further comprising ignoring the utterances that are longer than a specific threshold.
- 1 19. The NLU model of claim 17, wherein if the sample utterances are human-human
- 2 utterances, the method further comprises not labeling utterances longer then a threshold
- 3 amount.